ISSN 1561-5359. Штучний інтелект, 2018, № 1

УДК 004.93

О.С. Звенігородський[1], І.В. Качур[2]
[1]Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Україна
пр. Перемоги, 37, м. Київ, 03056
[2] Інститут проблем штучного інтелекту МОН і НАН України, Україна
пр. Академіка Глушкова, 40, м. Київ, 03680

МОДЕЛЬ СТРУКТУРИ СВІДОМОСТІ ЯК МНОЖИНИ ПРОЦЕСІВ

O.S. Zvenigorodsky[1], I.V. Kachur[2]
National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine
37, Peremogy avenue, Kyiv, 03056
[2]Institute of artificial intelligence problems MES and NAS of Ukraine, Ukraine
40, Academician Hlushkov av., Kyiv, 03680

THE CONSCIOUSNESS STRUCTURE MODEL AS A SET OF PROCESSES

Розглядається структура свідомості людини як множина перманентно діючих процесів, кожен з яких в основі представляється загальною структурою. Представлені складові загальної структури процесу. Обговорюється проблема штучної свідомості автономного робота.

Ключові слова: свідомість, процес, штучний інтелект

The structure of human consciousness is considered as a set of permanently operating processes, each of which based on a common structure. Presented components of the common structure of the process. The problem of artificial consciousness of an autonomous robot is discussed.

Key words: consciousness, process, artificial intelligence

Вступ

Питання моделювання інтелекту, свідомості і створення штучного інтелекту та штучної свідомості у світі й Україні обговорюються вже кілька десятиліть.

Зазначимо українську школу, яку створив академік М.М. Амосов. В Інституті кібернетики ім. В.М. Глушкова НАН України під керівництвом Палагіна О.В. [7] проводяться дослідження онтологічних моделей світу, методів видобутку знань, розробляються програмно-технічні засоби лінгвістичних процесорів для розуміння смислу текстів, що є складовими свідомості.

Проблема штучної (машинної, комп’ютерної) свідомості, ментальних станів, кваліа, мислення дискутується іноземними вченими протягом кількох десятиліть. Почали розвиватися логіко-філософські дослідження штучного інтелекту. Найбільш відомі праці J. Searle, M. Minsky, D. Dennet, R. Brooks, A, Sloman.

Дослідження цих вчених охоплюють спектр проблеми побудови комп’ютерної (машинної) свідомості. Обговорюються і пропонуються різни структури «свідомих» комп’ютерів, принципи і методологічні засади моделювання когнітивних, ментальних, емоційних і інших властивостей людини, що свідчить про актуальність проблеми і відсутність загальновизнаних теорій і концепцій у цій галузі.

Отримані результати показують, що проблема моделювання свідомості є актуальною, далека від вирішення і потребує деталізації та нових постановок задач.

Розглянемо основні підходи до даної проблематики. У [1] функції свідомості розглядаються з біологічної точки зору і представляються функціональними

моделями. Взагалі, свідомість представляється як рух активності у значимих моделях, посилений системою підсилення гальмування (СПГ), що відображає найважливіші відносини в системі суб'єкт-середовище. Суттєвим фактором такого підходу є те, що інтелект і моделі мають властивість-потребу, яка спонукає до дії. На вході потреби діє стимул, зміна якого призводить до задоволення потреби. Також зазначимо, що в даному підході мислення представляється як взаємодія моделей, що направляється почуттями і СПГ, а думка – модель, посилена СПГ в даний момент.

Автори J.A. Starzyk, D.K. Prasad [2] вважають, що машина є свідомою, якщо є механізми: сприйняття, дії, навчання, асоціативної пам'яті; центральний пристрій управління (ЦПУ), що керує усіма процесами свідомості – мотивацією, вибором мети, переключенням уваги, семантичною і епізодичною пам'яттю. ЦПУ використовує когнітивне сприйняття і когнітивне розуміння мотивацій, думок, або планів контролю навчання, уваги, мотивацій и контролю дій. Таким чином ЦПУ зв'язує когнітивний досвід з внутрішніми мотивами, чим створює самосвідомість і свідомий стан.

Привертають увагу аксіоми машинної свідомості [3]: присутність агента у світі, уявлення і звернення до минулого і поточного досвіду (зберігання подій у формі асоціативних згадувань і уміння їх викликати при запиті, формування можливих сценаріїв і прогнозування результатів сценаріїв на основі досвіду), увага, як засіб, що визначає зміст досвіду (агент взаємодіє з об'єктом або подією, на яку звернена увага і використовує сенсорний збір даних, моторні дії, воля (включає також мотивацію, здатність формувати цілі, діяти для досягнення цілей), емоції (уміння визначати природу переживань у своєму контексті і множині емоцій).

На базі окремого підходу створено систему LIDA [4], що удосконалюється протягом багатьох років і використовує частково символічний і частково connectionist підходи. До складу системи входять блоки: сприйняття (Perception), прийняття рішень (Workspace), короткострокової пам'яті (Episodic Memory), механізму свідомості (Consciousness Mechanism) на основі авторської теорії Global Workspace theory's (GWT), пам'яті дій (Procedural memory), вибору поведінки (Action selection). Характерною відзнакою даної системи є те, що вона складається з жорстко зв'язаних між собою блоків.

А.І. Шевченко [5] у своїх багаторічних дослідженнях представляє моделі інтелекту людини і штучної особистості, що включає штучну свідомість і штучний інтелект. Велика увага приділяється когнітивним процесам отримання і використання знань. Як і в попередній системі, свідомість має жорстку структуру та зв'язки між різними за смисловим змістом (моторними, психологічними, абстрактними, морально-етичними) блоками.

Узагальнюючи вищевикладене, розглянуте у працях [1-5], виділимо основні складові, що можна вважати необхідними при моделюванні свідомості. До них відносяться наступні моделі: сприйняття зовнішнього світу за допомогою сенсорів (органів чуття), пізнання світу за допомогою знань, навчання, логічного мислення і мови, взаємодії зі світом і собі подібними шляхом прийняття рішень з формулюванням і досягненням мети, мотивацій, психологічних станів.

Постановка задачі розробки структури свідомості

Розглянуті вище підходи і моделі спрямовані на моделювання інтелекту і свідомості агента (автономного робота), який діє у просторі і часі і реалізує раціональну поведінку, наприклад, у найпростішому випадку оминає перешкоди. Як

правило, вони представлені структурами у вигляді блоків або моделей. На наш погляд, головними недоліками цих підходів є жорсткі зв'язки моделей і блоків між собою, послідовність їх функціонування у часі, зведення функціонування до послідовних дій блоків, відсутність паралельності і невизначеність щодо складу інформації, яка передається між блоками. Це не узгоджується з даними нейрофізіологічних досліджень, які отримують при вивченні активності мозку.

Досягнення нейронаук, що застосовують сучасні методи, такі як електроенцефалографія, магнітно-резонансна томографія, позітронно-емісійна томографія, дозволили змінити уявлення про функціонування нейронів, механізми комунікації нейронів та міжнейронні взаємодії.

Виходячи із сучасних уявлень про організацію мозкових процесів, можна констатувати, що в корі головного мозку, яку пов'язують зі свідомістю та інтелектом, процеси передачі і обробки інформації відбуваються послідовно-паралельно, зі швидкою іррадіацією збудження між відповідними нейронними ансамблями. Інша важлива деталь – нейрони і клітини, що їх оточують, одного типу у всій корі, а результати їх діяльності, які пов'язані зі свідомістю та інтелектом, різні за змістом (смислом). Тобто абстрактні (математичні формули, поняття, речі суспільного буття) і чуттєві образи (зорові, слухові і ін.) різні в розумінні складових інтелекту і свідомості, але мають однакову нейрофізіологічну природу. Вважається, що образи є основним способом взаємодії нейронних ансамблів між собою.

Також, у розглянутих вище моделях не враховується дія гуморальних регуляторів діяльності мозку – біологічно активних речовин – гормонів, дія яких тривала за часом, наприклад, серотоніна (зміни концентрації якого напряму пов'язують з психоемоційним станом) та його попередника гормона – регулятора добових ритмів мелатоніна й інших.

Виходячи з цього, можна висунути гіпотезу про те, що різні за суттю процеси свідомості та інтелекту фізіологічно реалізуються однаково, а отже природа інформації на вході й виході теж є однаковою, і це – образи.

Таким чином, різні за суттю процеси свідомості та інтелекту можуть бути представлені загальною функціональною моделлю, а таке складне явище як свідомість – множиною паралельних процесів, які реагують не тільки на сенсорні й абстрактні образи, а й на вплив внутрішніх процесів, не пов'язаних з інформацією – «гормони», які гальмують або підсилюють процеси протягом часу, що на порядок перевищують час реакції на образи.

Проводячи аналогію зі свідомістю людини, з точки зори свідомості як стану, виділяють: сон, розглянутий як період відпочинку; пильнування, або активний стан. У активному стані процеси в мозку відбуваються паралельно, але координуються поточною думкою, яка постійно переключається спонтанно або при появі небезпечних і нових образів із процесів, що складають свідомість. У стані сну процеси з образами відбуваються, але з частковою координацією.

Таким чином, модель свідомості, з урахуванням зазначених станів, представимо як множину напівнезалежних процесів від процесу поточної думки з висуненням гіпотези, що внутрішня структура цих процесів однакова або має деякий набір загальних функцій.

Загальна структура процесу

Узагальнюючи склад блоків і моделей, розглянутих у [1-5], їх функцій, інформації на вході і виході, психологічних станів свідомості (прагнення отримання

задоволення) введемо поняття стану процесу і його рівноваги. Мета процесу – повернення в стан рівноваги, яка порушується появою нової інформації на вході. Це викликає дії процесу по розпізнаванню образу на вході. У разі неоднозначності розпізнавання (ймовірність, що образ належить класу образів процесу нижче деякого порогу, розпізнаються декілька образів) або появи нового образу процес має змогу впливати на органи тіла і мозку, які можуть збільшити інформативність образу на вході. Наприклад, включаються м'язи тіла і здійснюється скерування очей на об'єкт. Відповідно до цих вимог загальний елементарний процес представлений на рис. 1.

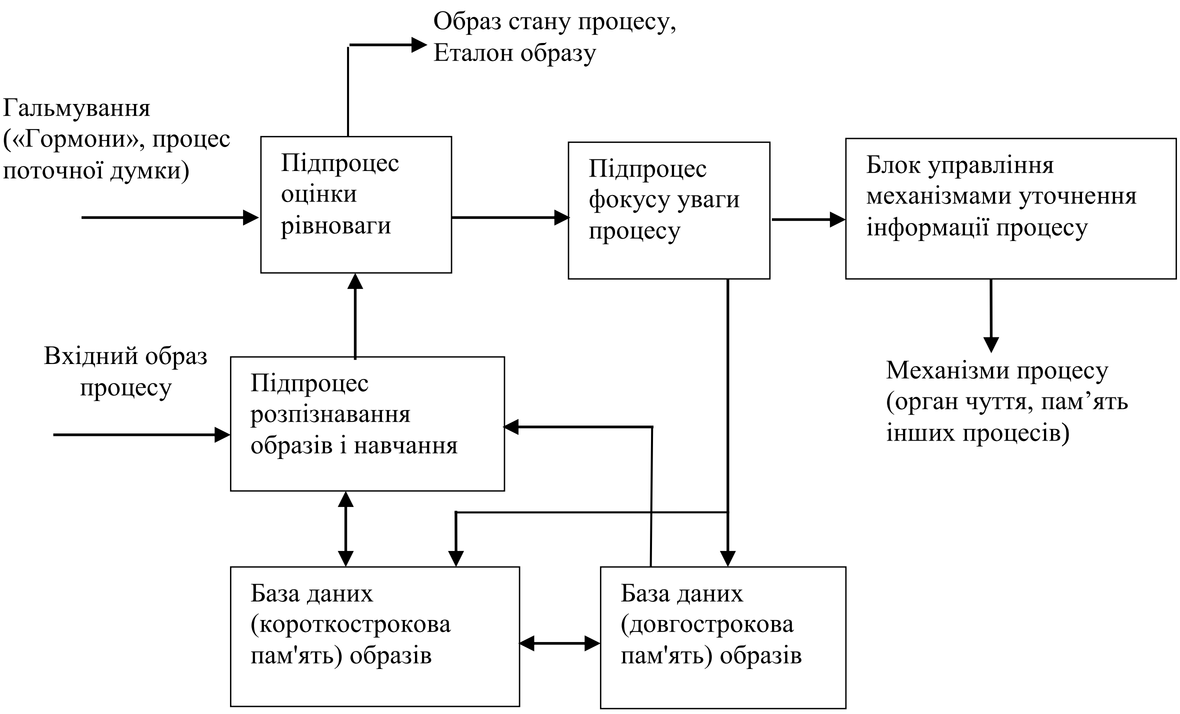


Рис. 1. Загальна структура процесу

До складу процесу входять наступні перманентні підпроцеси:

• Рівноваги. Рівновага, аналог задоволення із [1], як реакція на інформацію, що надходить у процес, визначається деяким рівнем збудженості процесу, який збільшується і процес виходить із рівноваги. Якщо образ розпізнається як такий, що відповідає процесу, то процес повертається у стан рівноваги, при цьому на вихід процесу надходить результат розпізнавання у вигляді образу. Вихідний образ складається з еталону розпізнаного образу і властивостей стану підпроцесу, що відображають, наприклад, рівень небезпеки образу для свідомості. Процес завжди працює, імітуючи процеси в нейронах, як біологічних клітинах, які зупиняються лише після смерті клітини. Можна зменшити активність процесу шляхом гальмування виходу, тоді на вихід проходять найбільш критичні для свідомості образи. Найпростіше загальмувати не всі його складові, а тільки частину, пов'язану зі станом рівноваги. По аналогії з людиною гальмування може бути двох типів: тривале, що викликається не інформаційними факторами, наприклад, збільшення концентрації «гормонів» і тимчасове гальмування процесу іншими процесами для

концентрації уваги свідомості на поточних думках. Крім цього, «гормони» можуть довготривало збуджувати процес в умовах відсутності образів на вході, звертаючи увагу інших процесів на себе, що можна розглядати як механізм мотивації у часі даного процесу.

• Розпізнавання і навчання. Головна функція підпроцесу – це розпізнавання образів і сприйняття нових образів (навчання). При розпізнаванні приналежність образу до класів образів процесу оцінюється деякою ймовірністю. Якщо ймовірність класифікації образу не перевищує деякий поріг, то включаються механізми фокусу уваги. Якщо образ класифікується як новий, то фокус уваги включає процеси навчання новому образу в короткостроковій і, за деяких умов, у довгостроковій пам'яті.

• Фокус уваги. Починає діяти в стані порушення рівноваги (невдоволення). Мета – збільшити інформативність образу для підвищення правильності розпізнавання. За аналогією з пригляданням, прислуховуванням, згадуванням і т.ін. При цьому на органи чуття і тіла надходять команди, що змінюють їх положення у просторі, призводять до зміни положення образу на рецептурному просторі сенсорів і зміни інформативності властивостей образу. За такою суттю діють і процеси з абстрактними образами, але вони активують і комбінують різні довгострокові пам'яті. Фокус уваги в кожному простому процесі пояснює здатність свідомості людини одночасно виконувати різні дії.

Функціонування підпроцесів забезпечують блоки і бази:

• Короткострокова пам'ять. Засіб забезпечення активності процесу розпізнавання протягом обмеженого часу і скорочення кількості образів у робочому просторі процесу. Аналог того, що одночасно людина сприймає не більше ніж 5 ± 2 об'єктів.

• Довгострокова пам'ять. Для швидкого розпізнавання (напівсвідомого) і переведення розпізнаного образу в короткострокову пам'ять.

• Блок управління механізмами уточнення інформації процесу за рахунок моторних дій з органами тіла і органів з мозком (відомо, що мозок складається з 10 відділів) для збільшення інформативності образів на вході і появи додаткової інформації від інших процесів.

Загальна структура свідомості

Процесів, що пов'язують зі свідомістю і інтелектом дуже багато. Виходячи з одного з визначень інтелекту, як розумної або раціональної поведінки, можна розглядати три класи: раціональна поведінка у просторі і часі, раціональна поведінка у соціумі, когнітивне сприйняття світу і свого Я. Найбільш важливі процеси для моделювання з цих класів ми виділили у групи. Загальна структура свідомості, що заснована на аналогіях зі свідомістю людини і висунутими гіпотезами, представлена на рис. 2.

Процес поточної думки – головний процес, що забезпечує функціонування свідомості в активному стані. Через нього проходять образи з усіх інших процесів (безперервні стрілки), він може гальмувати інші процеси (штрих пунктирні стрілки) і процеси можуть самостійно взаємодіяти між собою.



Рис. 2. Загальна структура свідомості

Процеси раціональної поведінки в часі і просторі відповідають за рух у просторі, який не шкодить фізичному тілу. Процеси логічно-асоціативного мислення зв'язують образи, що спостерігаються одночасно, для встановлення причинно-наслідкових зв'язків і прогнозування в майбутньому ситуацій при появі одного зі зв'язаних образів.

Процеси розуміння сенсу текстів забезпечують розуміння усної мови, письмових текстів, жестів, мови мистецтв. Процеси когнітивного пізнання відповідають за набуття знань про навколишній світ на побутовому і науковому рівні. Морально-етичні процеси і естетичні процеси відповідають за раціональну поведінку в соціумі.

Обговорення

Представлена структура свідомості людини і структура процесів, з яких складається свідомість, є спробою інтерпретації природи свідомості людини. На даний час у світі дискутабельною серед вчених різних галузей – футурологів, філософів, біологів – є проблематика штучної свідомості і штучного інтелекту роботів. Чи повинні вони мати всі властивості свідомості людини, відчувати як людина і бути повноцінними членами людського суспільства? З точки зору авторів – ні. Маючи знання про структуру свідомості, відповідні моделі процесів, уявлення про механізми процесів свідомості, робот буде виступати з позиції третьої особи. За наявності моделей цих процесів або доступу до них у мережах, якщо вони надто складні, інформації про людей, з якими спілкується, інформації про поточне зовнішнє середовище, своєї власної доповненої інформації із мереж, немає необхідності у власних морально-етичних та естетичних процесах, психоемоційних переживаннях та ін. Вищезазначене дозволить реалізовувати свою раціональну поведінку, яка буде сприйматись свідомою.

Можна дійти висновку, що штучний інтелект це є знання про процеси свідомості та інтелект людини. Якщо зважати на знання, накопичені людством, різноманіття культур, психологічних типів особистостей, то програмна модель робота здається буде дуже складною. Але аргументація щодо застосування сучасних мережевих технологій, що дозволяють моделі будь-якої складності розміщувати в «хмарі» і звертатись до них за необхідності, вирішує цю проблему.

Тобто, логічно можна дійти висновку, що мова вже йде не про штучний інтелект, а про розподілений штучний інтелект. Сучасні технології дозволяють створювати системи, що поєднують суперкомп'ютери, сервери і просто персональні комп'ютери в єдину мережу або множину мереж. Для автономних роботів можна закласти в даному роботі мінімальну кількість функцій, а складні моделі людської свідомості, до яких робот звертається за необхідності, можна передбачити за мережевим доступом. Перевага безсумнівна – ці моделі постійно удосконалюються та оновлюються.

Якщо відбудеться накопичення моделей властивостей і процесів свідомості та інших процесів інтелектуальної діяльності, можна створити на кшталт ноосфери В.І. Вернадського деяку розподілену ноосферу штучного інтелекту. Ноосферну концепцію штучного інтелекту з урахуванням трансдисциплінарної парадигми та з дослідженням методів і засобів інформаційно-технологічної підтримки ноосферогенези розвивають колектив авторів на чолі з О.В. Палагіним у [6].

Висновки

Запропонована єдина структура процесів свідомості може стати ефективним засобом практичного моделювання свідомості людини. Дана структура реалізується в парадигмі об'єктно-орієнтованого програмування із застосуванням сучасних методів розпізнавання, машинного і глибинного навчання.

Незважаючи на те, що необхідною умовою навіть індивідуальної свідомості людини є соціум, показано, що не всі складові свідомості людини необхідно моделювати у штучній свідомості автономного робота при включенні його в суспільне життя.

Представлення свідомості множиною процесів і дослідження гіпотез функціонування свідомості та інтелекту в такій моделі стане важливим фактором у постановці задачі визначення функцій і створення штучної свідомості і штучного інтелекту автономних роботів і глобального розподіленого штучного інтелекту.

Література

1. Амосов Н.В. Алгоритмы разума. Киев – Издательство «Наукова думка», 1979, 224 с.
2. Starzyk J.A., Prasad D.K. A computational model of machine consciousness International Journal of Machine Consciousness. - 2014. – Vol. 3 (2), P. 255-282. https://doi.org/10.1142/S1793843011000819
3. Aleksander I. The potential impact of machine consciousness in science and engineering// International Journal of Machine Consciousness, 1(1), 2009, pages 1-9.
4. Franklin S., Mad T., D'Mello S., Snaide J. LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning, IEEE Transactions on Autonomous Mental Development, Volume: 6, Issue: 1, March 2014. Pages 19–41, doi: 10.1109/TAMD.2013.2277589
5. Шевченко А.И. К вопросу о создании искусственного интеллекта / А.И. Шевченко // Штучний інтелект. – 2016. – № 1. – С. 7-15.
6. Palagin, A.V., Kurgaev, A.F., Shevchenko, A.I. The Noosphere Paradigm of the Development of Science and Artificial Intelligence Cybernetics and Systems Analysis, Vol. 53, No. 4, July, 2017 Pages 503-511, doi: 10.1007/s10559-017-9952-4

Literatura

1. Amosov N.V. Algoritmy razuma Kiyev – Izdatelstvo «Naukova dumka». 1979. 224 s.
2. Starzyk J.A., Prasad D.K. A computational model of machine consciousness International Journal of Machine Consciousness. - 2014. – Vol. 3 (2), P. 255-282. https://doi.org/10.1142/S1793843011000819
3. Aleksander I. The potential impact of machine consciousness in science and engineering// International Journal of Machine Consciousness, 1(1), 2009, pages 1-9.
4. Franklin S., Mad T., D'Mello S., Snaide J. LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning, IEEE Transactions on Autonomous Mental Development, Volume: 6, Issue: 1, March 2014. Pages 19–41, doi: 10.1109/TAMD.2013.2277589
5. Shevchenko A.I. K voprosu o sozdanii iskusstvennogo intellekta / A.I. Shevchenko // Shtuchniy intelekt. – 2016. – № 1. – S. 7-15.
6. Palagin, A.V., Kurgaev, A.F., Shevchenko, A.I. The Noosphere Paradigm of the Development of Science and Artificial Intelligence Cybernetics and Systems Analysis, Vol. 53, No. 4, July, 2017 Pages 503-511, doi: 10.1007/s10559-017-9952-4

RESUME

O.S. Zvenigorodsky, I.V. Kachur

The consciousness structure model as a set of processes

Summarizing the developments in the field of creating artificial consciousness, the main components that are modeled are distinguished. Based on the data from the physiology of neural processes of the brain, a hypothesis is advanced that different in the sense processes related to consciousness functionally are realized in the same way, because the processes of consciousness occur in the cerebral cortex, and the neurons that realize them do not differ from each other. Based on this, the structure of the process is proposed, in which the input and output information are images. The main function of the process is pattern recognition. The recognition result determines the stability of the process. If the recognition result is positive, the process returns to a stable state. In the process of recognition, are involved the focus attention process, a short-term and long-term memory of images and instruments that help to obtain additional information (focusing).

A general structure of consciousness is proposed in the form of a set of permanently operating processes. The main groups of these processes are distinguished. In the center of the structure is the "Process of current thought", which strengthens or inhibits the activity of all other processes. In this way, the ability of a person to switch attention between thoughts is modeled. In this case, the semantic processes do not stop interaction with each other. Such a structure is convenient for carrying out research and obtaining knowledge about the principles of the functioning of human consciousness and intelligence.

The issues of intellectuality of robots and distributed artificial intelligence are discussed in the case of using this structure, as well as issues of practical implementation.

Надійшла до редакції 15.01.2018

